The Behavior Analyst 2013, 36, 137–143 No. 1 (Spring)

Go Forth and Be Variable

Gregory J. Madden

In the target article, Vyse's (2013) core argument is that the experimental analysis of behavior (EAB) has painted itself into a methodological corner. It is time to retire, or at least curtail the use of, the operant chamber because it has already yielded its most important discoveries, and the costs of further discoveries outweigh their probable benefits. Vyse suggests that much contemporary research in EAB is well controlled but esoteric, arguing about little, and with little prospect for affecting the world at large. Given this perceived state of affairs, Vyse makes recommendations for change; paramount among these is to expand the methods (research questions asked, procedures and apparatus used, research designs, and data-analysis techniques) that have defined the field of behavior analysis. So as to begin on a positive note, I will tackle these arguments in reverse order.

The Merits of Variability

If scientific practice is conceptualized as an evolving unit within an iterative selection-based system, then instances of variability are critical to the forward movement of the cultural practice. Without variability, units will survive if they are compatible with the current contingencies of selection, but they will fail to survive when those contingencies change. Thus, as change comes to the contingency space in which behavior analysis must survive, there is merit in Vyse's call for variability. Many a company within a selection-based capitalist economy failed to survive when it behaved with stereotypy despite evolving consumer demands. Of course, most start-up companies and biological mutations fail and, analogously, most researchers who embark on a truly innovative research program will experience failure. As notable scientists such as Skinner (1956), Kahneman (2011), and now Vyse have put it, there is a good deal of luck in any successful research endeavor, and those who behave most variably will find that there are plenty of turning points at which the research can fail. Nonetheless, scientific progress is built on the backs of those who failed, with history remembering only those lucky few who innovated and succeeded.

The variability that Vyse argues for need not be a risky research line. Instead, he suggests some less extreme instances of variability that might be profitably adopted. Prominent among these is his suggestion that behavior analysts should adopt more flexibility in their research methods and data-analytic strategies. I agree with this suggestion. As Vyse implies, research methods are tools, and no single tool is appropriate for every job. For example, in my own laboratory my graduate students and I use a variety of research designs because not every question can be answered easily with the same design. For example, a single-subject design is not appropriate if one is evaluating the effects of a protracted training history versus no such history on rats' delay discounting and drug self-administration (Stein et al., 2013). Recognizing this, and embracing variability in research design, the current and past two editors of the *Journal of the Experimental Analysis of Behavior* (*JEAB*) have explicitly taken a utilitarian and welcoming

Correspondence concerning this article should be addressed to Gregory J. Madden, Department of Psychology, 2810 Old Main Hill, Utah State University, Logan, Utah 84322 (e-mail: greg.madden@usu.edu).

approach to reports that use designs other than single-subject research designs (Green, 2004; Madden, 2012; Mazur, 2008).

Second, to execute the kind of systematic and expansive basic research program that is most likely to contribute to knowledge and future technologies, one needs substantial research funding. This has always been true, but the contingencies in academia have changed in the last 20 years such that extramural funding is a requirement for tenure at many Carnegie Classified Research 1 universities. Given tight state budgets and concerns about continuously rising tuition expenses, I think the days of tenure without extramural grant funding is going the way of the dinosaur. Responding with intelligent variability to this contingency change will require the kind of graduate training discussed by Vyse. Behavior analysts who submit grant proposals containing single-subject research designs coupled with visual inspection of data will, in my experience, find that their proposals are not funded by the National Institutes of Health or the National Science Foundation. The prospects of basic behavior-analytic research are grim if our PhD programs do not prepare their graduates to write competitive grants that include sophisticated statistical analyses. Some may argue that this is tantamount to selling one's soul, but I have always found compelling an argument made in this journal by Michael Davison (1999): One can still conduct well-controlled steady-state single-subject research while using inferential statistics. Doing so requires increasing your sample size to at least six subjects, but in doing so the researcher will convince behavior analysts *and* those whose training has prepared them to understand the importance of inferential statistics and to regard visual inspection as an unacceptable practice.

And it is not just basic researchers. Applied behavior-analytic researchers are hamstrung if they do not have the training necessary to understand and skeptically evaluate research published outside traditional behavior-analytic journals. Having assigned some of the latter category of articles to graduate students who were not required to take courses in traditional research designs and inferential statistics, I was surprised to find that these students often either (a) rejected the findings in a knee-jerk fashion, without an ability to craft a reasoned argument against the paper; or (b) accepted the findings whole cloth, with no ability to identify the weaknesses of the paper and, as a result, no ability to design a follow-up study to further evaluate the claims of the paper's authors. Either outcome is unacceptable. If (a) is true, then the student has dismissed a potentially useful independent variable that might be integrated into a behavioral intervention. If (b) is true, the student might integrate a useless independent variable into their behavior-change endeavors. By understanding research designs and inferential statistics well enough to evaluate when they have been properly and improperly used, the behavior analyst will be in a position to benefit from the work of a larger body of social scientists and will be able to mount a convincing critique when a study does not deliver what it claims.

In supporting Vyse's call for an expansion of our data-analytic repertoire, I recommend that behavior analysts follow closely the literature on inferential statistics that are being developed to analyze short, autocorrelated data streams. For example, a potentially useful procedure is the Simulation Modeling Analysis (SMA; Nash, Borckardt, Abbasa, & Gray, 2011). The SMA uses bootstrapping techniques to randomly generate 5,000 time-series data sets that contain the same autocorrelation as the obtained data. When one of these data sets more closely approximates a theoretical intervention-phase vector (i.e., what an intervention might optimally produce) than do the original data, the data set is flagged and counted to estimate a probability of randomly obtaining an intervention data stream with an effect as large as was empirically obtained (for a review of this and other inferential statistics being developed for single-subject designs, see Borckardt, Nash, Balliet, Galloway, & Madan, 2013). For those who wish to explore the capability of the SMA, the software used to conduct this analysis is freely available at http://clinicalresearcher.org/software.htm. The SMA, and the next generation of time-series inferential statistics, may offer behavior analysts an avenue for communicating with non-behavior analysts (e.g., most members of extramural grant-review panels, who regard $p < .05$ to be a critically important criterion).

A related set of Vyse's prescriptions for variability is that behavior analysts should read more widely, should be more open to self-report data, and should acquire the quantitative skills needed to talk to psychologists, economists, and presumably other scientific disciplines. By doing so, behavior analysts prepare themselves to recognize the importance of their own and others' basic research findings in addressing societal problems that involve human behavior. Vyse's example of delay discounting is one with which I can identify, and I hope the reader will forgive me for the nostalgic recollections that follow. I offer them as a case study of sorts.

In 1996, Warren Bickel, Nancy Petry, and I were discussing an article on delay discounting that was published in *JEAB* by Green and Myerson (1995). Inspired by the paper, we decided to undertake a study to see if opioid-dependent individuals more steeply discounted delayed rewards than did a matched control group. The study confirmed our intuition, and we published the findings in an American Psychological Association journal (Madden, Petry, Badger, & Bickel, 1997). A follow-up study undertaken in the Bickel lab demonstrated that cigarette smokers also more steeply discounted delayed rewards than did a matched control group (Bickel, Odum, & Madden, 1999). These findings were published in a prominent journal in the psychopharmacology literature.

This illustrates a case study in which the researchers followed several of the practices that Vyse argues for in his target article and, perhaps as a result (and with a bit of luck), experienced success. First, Bickel's lab group was reading widely and was, therefore, prepared to recognize the potential importance of an interesting *JEAB* article for expanding our understanding of substance abuse. Second, we published our findings in traditional journals so that they could reach a large audience of substance-abuse researchers. Perhaps as a result of this decision, the two articles referenced above have been cited more than 1,000 times (combined). Third, the nature of the question that we asked required a between-groups design and inferential statistics. Using these methods may have made it easier to publish our work in journals that would be more likely to reach our target audience. Fourth, consistent with Vyse's advice, we used self-report measures of choice; participants considered two hypothetical monetary consequences and were instructed to make their decisions as though the outcomes were real. I would say that such methods would never be published in *JEAB* except that the Green and Myerson (1995) paper (published in *JEAB*) that inspired our work was a study of self-reported choices that involved hypothetical rewards. Indeed the methods we employed in our delay discounting studies were almost entirely borrowed from a *JEAB* article published by Rachlin, Raineri, and Cross (1991). I admit that self-report measures are a minority in the pages of *JEAB,* but other well-cited examples could be identified. My final point, which has little to do with the target article but cannot be resisted, is "When you run onto something interesting, drop everything else and study it" (Skinner, 1956, p. 223). Those I was lucky enough to work with in the Bickel lab in the mid-1990s followed this advice and have made the sort of impact on psychology, economics, and neuroscience that Vyse envisions (e.g., Bickel, Miller, Kowal, Lindquist, & Pitcock, 2007; Koffarnus, Jarmolowicz, Mueller, & Bickel, 2013; Odum, 2011; Petry, 2012). In sum, I concur with Vyse's arguments that behavior analysts should exhibit variability in the journals they read and those in which they publish. We should regard research methods as tools and be prepared to use the full toolbox so that we may answer every interesting question. Finally, we should take a pragmatic approach to our data-analysis techniques, and should modify our research practices (e.g., using enough subjects to allow the use of nonparametric inferential statistics; Davison, 1999) so that we will be prepared to use data analysis techniques that are viewed as valid by a wide audience.

The Costs and Benefits of EAB

And now I will briefly address the portion of Vyse's paper that I disagree with. According to Vyse, the most important principles that can be discovered in operant conditioning chambers have been discovered, and the costs of extracting further principles, or refinements of existing principles, outweigh the potential benefits. I will offer four somewhat off-the-cuff defenses of contemporary EAB.

First, defending contemporary research in EAB is as difficult as criticizing it; one cannot predict the future, so it is impossible to know which piece of "esoterica," if any, will prove useful in improving our understanding of and ability to influence behavior. Consistent with this, I feel certain that when the first studies of human response patterns under fixed-interval schedules of reinforcement were published (e.g., Laties & Weiss, 1963; Lipman & Meyer, 1967), the authors would not have predicted that their findings would launch an extensive empirical study of human rule-governed behavior and variables that underlie sensitivity of behavior to changing contingencies of reinforcement (e.g., Galizio, 1979; Hayes, Brownstein, Zettle, Rosenfarb, & Korn, 1986). Likewise, these researchers from the 1960s would not have guessed that this extended research line would subsequently be combined with findings in the stimulus equivalence (and other relational responding) literature, to influence the writing of those who would affect the direction of contemporary clinical psychology (e.g., Hayes, Strosahl, & Wilson, 1999, 2012). Standing in 2013 and looking on contemporary EAB research, one might question the utility of seemingly esoteric research on, for example, the acquisition of stimulus–stimulus relations by nonhumans (e.g., Urcuioli, Jones, & Lionello-DeNolf, 2013). However, when one closely reads and considers the models presented in these papers, one can see how they advance a fundamental understanding of relating behavior. Predicting the future is difficult, but as I read papers like Urcuioli et al. (2013), I do not have to stretch my imagination too much to consider their translational potential in an area as important as teaching humans to read (Saunders, 2011).

Second, you know that sense that most of us (in the United States) have that Saturday Night Live (SNL) is not as funny as it used to be? I think something similar occurs when we reminisce about the good old days of

EAB. We tend to remember the funniest of the SNL skits and forget that most of the skits that were broadcast live just weren't that funny. The latter skits are those that we have forgotten, but if you go back and watch an entire repeat broadcast of SNL from whatever era you consider to be the "golden age of SNL," you will find that most of the skits don't elicit even a chuckle. Likewise, with basic research in EAB, some studies published each year will have immediate impact, some will have an impact at a later date, and some just never will be "funny." I recall that after I had read many of the most important studies in *JEAB*, it seemed to me that the new issues of *JEAB* just were not up to the standards of the past. I suspect this is because I had, until that time, been reading *JEAB*'s "greatest hits." This is analogous to saying that contemporary SNL broadcasts are not as funny as the "Best of SNL" video that you just watched. In an iterative selection-based system, not every instance of variability will be selected; some of the SNL skits broadcast this week will seem esoteric; but years from now, when we can recall only the funny skits … ah, the golden age.

Third, and quite contrary to Vyse's assessment, I actually think that some of the contemporary research conducted in operant chambers in EAB is ground breaking and could, 20 years from now, prove to be more important than the research conducted in the golden age of EAB. I cannot possibly list every instance of this ground-breaking research but, beyond the Urcuioli et al. (2013) paper mentioned above (the latest in this important research line being conducted in the Urcuioli lab), I would point to the ongoing empirical debates about the principle at the very core of behavior analysis: reinforcement. For those who are unfamiliar with these debates, a number of researchers have presented findings in behavior-analytic and other journals suggesting that reinforcers (primary and secondary) do not strengthen or increase the probability of the behavior that they follow (e.g., Davison & Baum, 2006; Krägeloh, Davison, & Elliffe, 2005; Mazur & Biondi, 2013; Shahan & Podlesnik, 2008). Instead, these consequent stimuli provide "signpost" information about, for example, the time to the next occurrence of a phylogenetically important event (e.g., the delivery of food). Other researchers are skeptical of this discriminative-control position and conduct important research that explores the assumptions of old and new accounts of this process (e.g., Bell & Williams, 2013). The debate continues and, in my opinion, it is not an esoteric debate; if we better understood the "reinforcement" process, the translational research that followed would, I presume, be important indeed.

Finally, Vyse sees "considerable evidence of advances in other areas of psychology" (p. 124). As should be clear by this point, I agree with this statement. However, it has always struck me that the allocation of human resources in psychology among the various schools of thought (e.g., perhaps 30% of today's graduate students will study cognitive psychology, another 20% might study child development, 5% will study behavior analysis, etc.) does not result in proportional evidence-based behavior or cognition-influencing technologies from each school of thought (e.g., 5% of the evidence-based technologies come from behavior-analytic research). I don't know how our human resources are being allocated these days, but it is a safe bet that there are far more neuroscientists and cognitive psychologists than there are behavior analysts. Rarely do we get an accounting of the return on our human-resource investments, and it would be difficult to devise a way to evaluate this that would be acceptable to all. One approximation, however, might be the 2012 report of the Inter-Organizational Task Force on Cognitive and Behavioral Psychology Doctoral Education (Klepac et al., 2012). This task force was led by the Association for Behavioral and Cognitive Therapies, a member organization of primarily clinical researchers. The task force's mission was to recommend competencies that should be held by those who graduate from doctoral programs that train clinical, counseling, and school psychologists in cognitive behavioral therapies. Consistent with my sense that behavior analysts and behavior-analytic clinical psychologists have produced a disproportionate return on investment, at least 66% of the procedures in the task force's list of evidence-based interventions were behavior analytic in nature (e.g., shaping, self-management and habit reversal, extinction or exposure, behavioral activation). In support of my earlier claim that clinical behavior analysts are influencing the direction of contemporary clinical research and practice, knowledge of core concepts in acceptance and commitment therapy (Hayes et al., 2012) were included in this list of evidence-based intervention competencies (e.g., defusion, values clarification).

To be sure, the task force report lists many more clinical skills competencies that behavior analysts have, to the best of my knowledge, not contributed to (e.g., cognitive neuroscience), but I believe that behavior analysis has been a good return on investment. Given the vibrancy of the empirical debate that continues in EAB and the peppering of funny skits that may still be found on SNL, I must disagree with Vyse's assessment that the future costs of EAB are likely to outweigh the benefits produced. When choosing a stock in which to invest, one looks for a company with a long track record of success. Behavior analysis has a great track record, but so did the railroads. Therefore, I close by underscoring my support for Vyse's calls for intelligent variability in the training and research practices of behavior analysts. The contingencies of selection have changed and will change again. Go forth and be variable.

REFERENCES

Bell, M. C., & Williams, B. A. (2013). Conditioned reinforcement in chain schedules when time to reinforcement is held constant. *Journal of the Experimental Analysis of Behavior*, *99*, 179–188.

Bickel, W. K., Miller, M. L., Kowal, B. P., Lindquist, D. M., & Pitcock, J. A. (2007). Behavioral and neuroeconomics of drug addiction: Competing neural systems and temporal discounting processes. *Drug and Alcohol Dependence*, *90, Supplement 1*, S85–S91.

Bickel, W. K., Odum, A. L., & Madden, G. J. (1999). Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. *Psychopharmacology*, *146*, 447–454.

Borckardt, J. J., Nash, M. R., Balliet, W., Galloway, S., & Madan, A. (2013). Time-series statistical analysis of single-case data. In G. J. Madden (Ed.), *Handbook of behavior analysis: Volume 1. Methods and principles* (pp. 251–266). Washington, DC: American Psychological Association.

Davison, M. (1999). Statistical inference in behavior analysis: Having my cake and eating it? *The Behavior Analyst*, *22*, 99–103.

Davison, M., & Baum, W. M. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior*, *86*, 269–283.

Galizio, M. (1979). Contingency-shaped and rule-governed behavior: Instructional control of human loss avoidance. *Journal of the Experimental Analysis of Behavior*, *31*, 53–70.

Green, L. (2004). Editorial. *Journal of the Experimental Analysis of Behavior*, *81*, 1–3.

Green, L., & Myerson, J. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, *64*, 263–276.

Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I., & Korn, Z. (1986). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior*, *45*, 237–256.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: Guilford.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). *Acceptance and commitment therapy:*

The process and practice of mindful change (2nd ed.). New York, NY: Guilford.

Kahneman, D. (2011). *Thinking fast and slow*. New York, NY: Farrar, Straus and Giroux.

Klepac, R. K., Ronan, G. F., Andrasik, F., Arnold, K. D., Belar, C. D., Berry, S. L., … Dowd, E. T. (2012). Guidelines for cognitive behavioral training within doctoral programs in the United States: Report of the Inter-Organizational Task Force on Cognitive and Behavioral Psychology Doctoral Education. *Behavior Therapy*, *43*, 687–697.

Koffarnus, M. N., Jarmolowicz, D. P., Mueller, E. T., & Bickel, W. K. (2013). Changing delay discounting in the light of the competing neurobehavioral decision systems theory: A review. *Journal of the Experimental Analysis of Behavior*, *99*, 32–57.

Krägeloh, C. U., Davison, M., & Elliffe, D. M. (2005). Local preference in concurrent schedules: The effects of reinforcer sequences. *Journal of the Experimental Analysis of Behavior*, *84*, 37–64.

Laties, V. G., & Weiss, B. (1963). Effects of a concurrent task on fixed-interval responding in humans. *Journal of the Experimental Analysis of Behavior*, *6*, 431–436.

Lipman, L. G., & Meyer, M. E. (1967). Fixed interval performance as related to instructions and to subject's verbalizations of the contingency. *Psychonomic Science*, *8*, 135–136.

Madden, G. J. (2012). Editorial: Lessons from *JEAB* to *JEAB*. *Journal of the Experimental Analysis of Behavior*, *97*, 1–4.

Madden, G. J., Petry, N. M., Badger, G. J., & Bickel, W. K. (1997). Impulsive and self-control choices in opiate-dependent patients and non-drug-using control participants: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology*, *5*, 256–262.

Mazur, J. E. (2008). Editorial: Half a century. *Journal of the Experimental Analysis of Behavior*, *89*, 1–3.

Mazur, J. E., & Biondi, D. R. (2013). Pigeons' choices with token stimuli in concurrent variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, *99*, 159–178.

Nash, M. R., Borckardt, J. J., Abbasa, A., & Gray, E. (2011). How to conduct and statistically analyze case-based time series studies, one patient at a time. *Journal of Experimental Psychopathology*, *2*, 139–169.

Odum, A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior*, *96*, 427–439.

Petry, N. M. (2012). Discounting of probabilistic rewards is associated with gambling abstinence in treatment-seeking pathological gamblers. *Journal of Abnormal Psychology*, *121*, 151–159.

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, *55*, 233–244.

Saunders, K. (2011). Designing instructional programming for early reading skills. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 92–109). New York, NY: Guilford.

Shahan, T. S., & Podlesnik, C. A. (2008). Conditioned reinforcement value and resistance to change. *Journal of the Experimental Analysis of Behavior*, *89*, 263–298.

Skinner, B. F. (1956). A case history in scientific method. *American Psychologist*, *11*, 221–233.

Stein, J. S., Johnson, P. S., Smits, R. R., Renda, R., Liston, K. J., Shahan, T. S., & Madden, G. J. (2013). Early and prolonged exposure to reward delay: Effects on impulsive choice and alcohol self-administration. *Experimental and Clinical Psychopharmacology.* Advance online publication.

Urcuioli, P. J., Jones, B. M., & Lionello-DeNolf, K. M. (2013). Evidence for response membership in stimulus classes by pigeons. *Journal of the Experimental Analysis of Behavior*, *99*, 129–149.

Vyse, S. (2013). Changing course. *The Behavior Analyst*, *36*, 123–135.